

修 士 論 文 の 和 文 要 旨

研究科・専攻	大学院 電気通信学 学研究科 情報通信工学 専攻 博士前期課程		
氏 名	鶴原 翔夢	学籍番号	0630046
論文題 目	統計的手法を用いた Web 閲覧履歴の解析と Web 推薦への応用		
<p>要 旨</p> <p>Web 上に存在する情報量が肥大・多様化し、情報爆発と言われる現象が発生している。こうした状況下では情報量があまりに多すぎるために、個人にとって有用な情報を得ることが難しくなっている。これを改善するべく、膨大なWeb ページの中から有用なものだけを選択し、提示してくれる Web ページ推薦システムが研究されている。</p> <p>本研究では、ユーザのWeb ページ閲覧履歴を収集し、これを解析することでユーザにとって有用なページを推薦する。推薦のアルゴリズムとしては、協調フィルタリングを用いる。ユーザの閲覧履歴を比較し、似た閲覧履歴を持っているユーザを見つけ出し、そのユーザの履歴を元に推薦を生成するものである。</p> <p>従来の研究では、ユーザの閲覧履歴が似ているかどうかを示す指標である類似度の計算法として、同一のページを閲覧しているかどうか、など単純な尺度に基づいた手法を用いていた。膨大な量のページが存在している中で、異なる二人のユーザが同一のページを閲覧していることはまれであり、このような類似度計算法では多くのユーザ間の類似度が 0 となってしまう推薦が失敗するケースが多いことが問題となっていた。本研究は、収集したユーザの閲覧履歴に対して潜在意味解析や独立成分分析といった統計的手法を適用することで、従来法の問題を解決することを目的とする。統計解析の結果、閲覧履歴データ空間上の特徴軸が得られる。本論文では、この特徴軸を閲覧コンセプトと呼ぶ。また、その結果を使ってユーザ間の類似度を計算することを提案する。</p> <p>本学情報基盤センターより提供を受けた、電気通信大学のHTTP 対外接続ログデータを用いて実際に閲覧コンセプトを抽出した。また、これを用いた類似度の性質について、ソーシャルブックマークデータを用いた比較検討実験を行った。</p> <p>実際に被験者 8 名に対して Web ページ推薦を行った。比較のため、従来手法に基づいた推薦と統計処理を利用した推薦をランダムに混ぜて提示し、興味に沿う内容であったか、面白かったかという基準で評価を行ってもらった。その結果、従来手法による推薦に対する評価と統計処理を用いた推薦に対する評価の間に有意な差を得ることはできなかったが、統計手法を用いた推薦が有効な場合があることを示した。</p>			